Curve Your Enthusiasm: Concurvity Regularization in Differentiable Generalized Additive Models

Julien Siems*

Joint work with: Konstantin Ditschuneit*, Winfried Ripken*, Alma Lindborg*, Maximilian Schambach, Johannes Otterbach and Martin Genzel

*equal contribution

1. Generalized Additive Models

Model

➔ **Linear Regression:**

$$f(\mathbf{x}) = \beta + w_1 x_1 + ... + w_n x_n$$

➔ **Generalized Additive Models (GAM):**

$$f(\mathbf{x}) = \beta + f_1(x_1) + \ldots + f_n(x_n)$$

In the past:
Splines

Neural Additive Models (NAM):

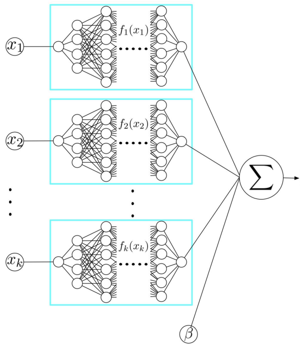
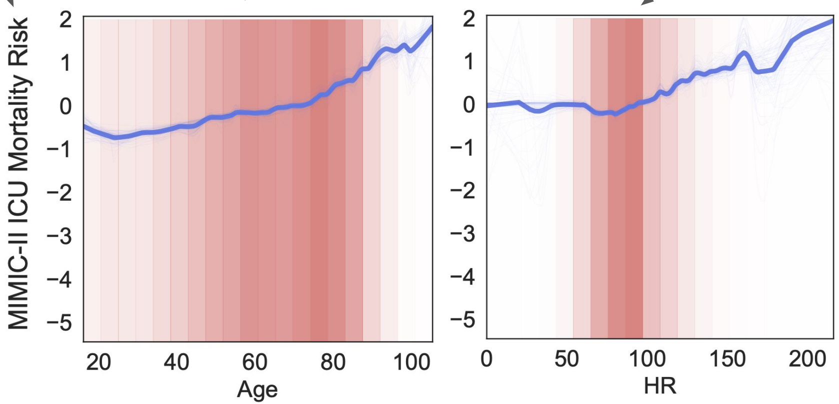


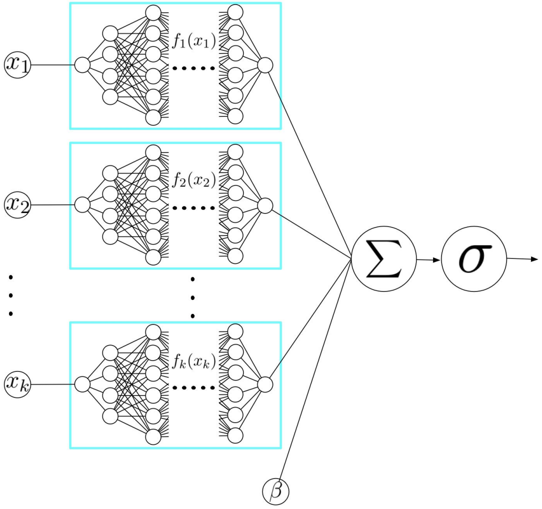
Figure from: Agarwal, Rishabh, Levi Melnick, Nicholas Frosst, Xuezhou Zhang, Ben Lengerich, Rich Caruana, and Geoffrey E. Hinton. "Neural additive models: Interpretable machine learning with neural nets." *Advances in Neural Information Processing Systems* 34 (2021): 4699-4711.

2. Generalized Additive Models

Neural Additive Models (NAM):

$$f(\mathbf{x}) = \beta + f_1(x_1) + \ldots + f_n(x_n)$$

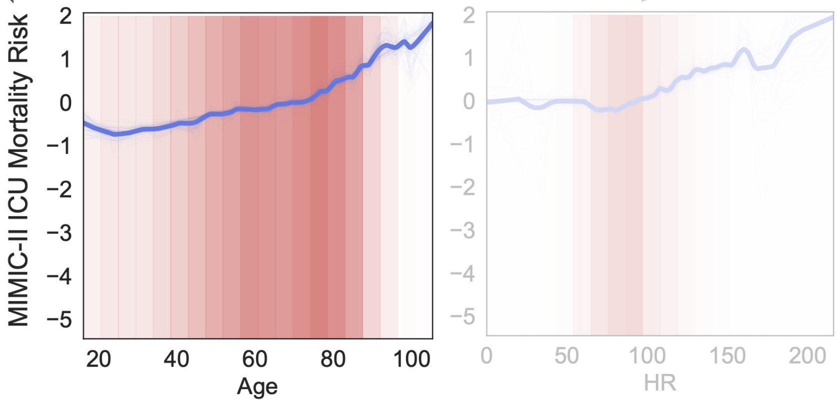


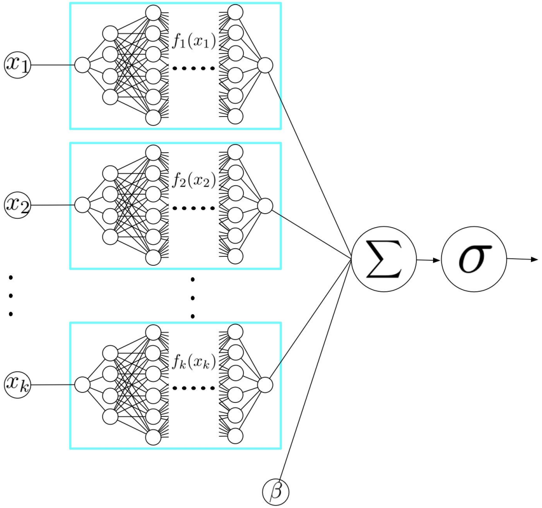


2. Generalized Additive Models

Neural Additive Models (NAM):

$$f(\mathbf{x}) = \beta + \underbrace{f_1(x_1)} + \ldots + \underbrace{f_n(x_n)}$$





Hypothesis: *GAMs allow to understand each features' contribution in* ***isolation****.*

3. Multicollinearity and Concurvity

Model

➔ **Linear Models:**

$$f(\mathbf{x}) = \beta + w_1 x_1 + ... + w_n x_n$$

➔ **Generalized Additive Models (GAM):**

$$f(\mathbf{x}) = \beta + f_1(x_1) + \ldots + f_n(x_n)$$

Problems

➔ **Multicollinearity**

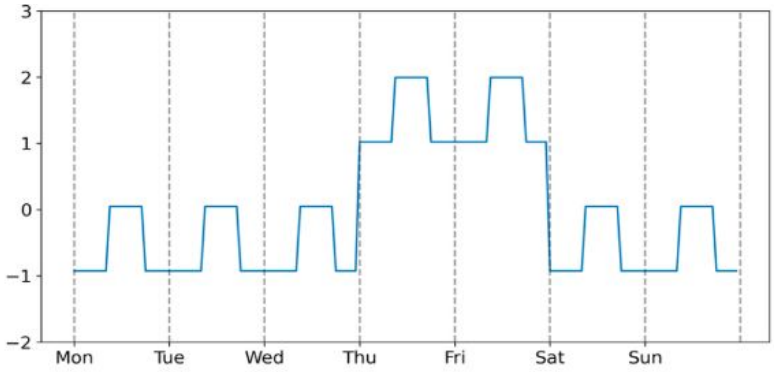
➔ **Concurvity**

3. Multicollinearity and Concurvity

Generalized Additive Model (GAM) $$f(\mathbf{x}) = \beta + f_1(x_1) + \ldots + f_n(x_n)$$

➔ **Concurvity**
Correlation between transformed features $f_1(x_1)$, ..., $f_n(x_n)$.

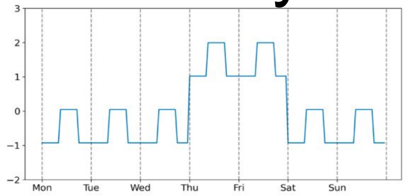
Example:



$$f(t) = S_{24h}(t) + S_{7d}(t)$$

3. Multicollinearity and Concurvity

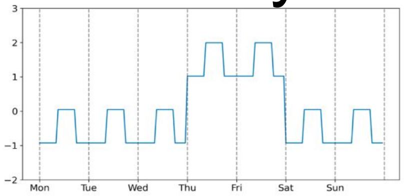
Task: Fit periodic timeseries with daily and weekly seasonality



$$f(t) = S_{24h}(t) + S_{7d}(t)$$

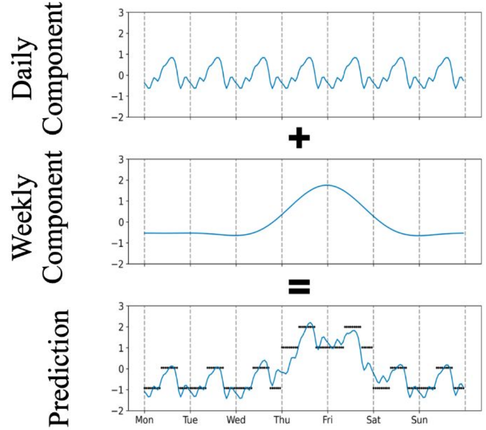
3. Multicollinearity and Concurvity

Task: Fit periodic timeseries with daily and weekly seasonality



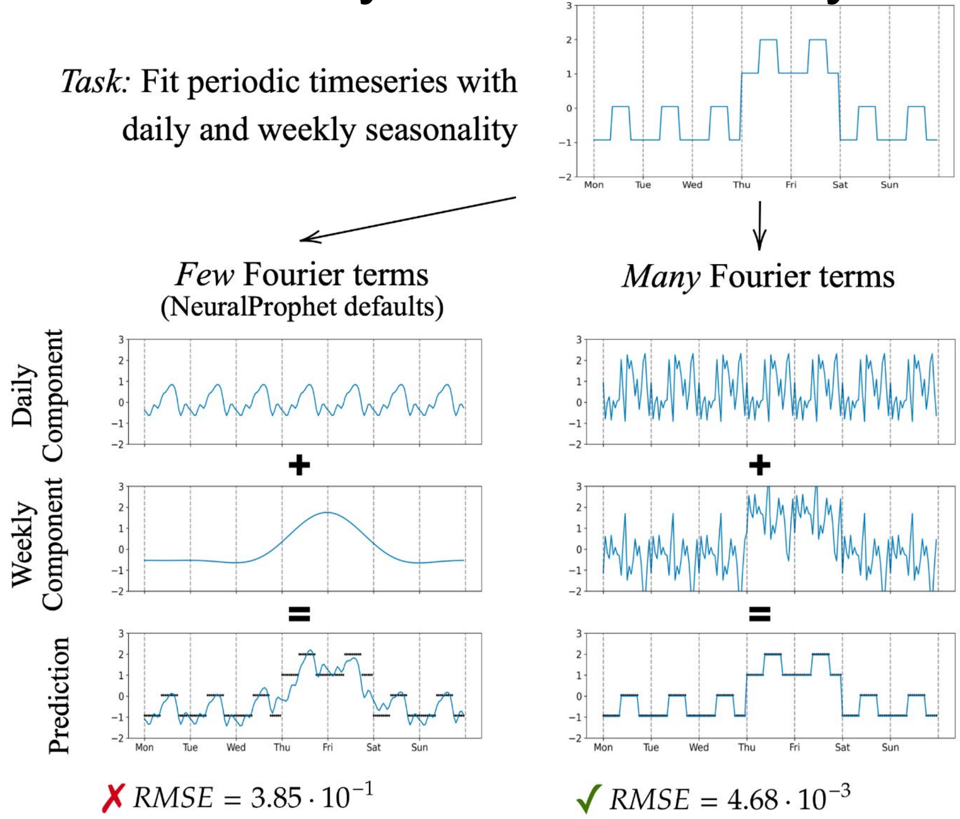
$$f(t) = S_{24h}(t) + S_{7d}(t)$$

Few Fourier terms
(NeuralProphet defaults)



✗ $RMSE = 3.85 \cdot 10^{-1}$

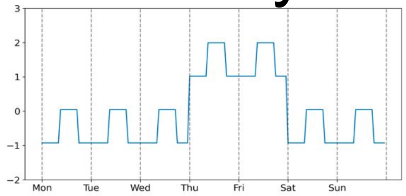
3. Multicollinearity and Concurvity



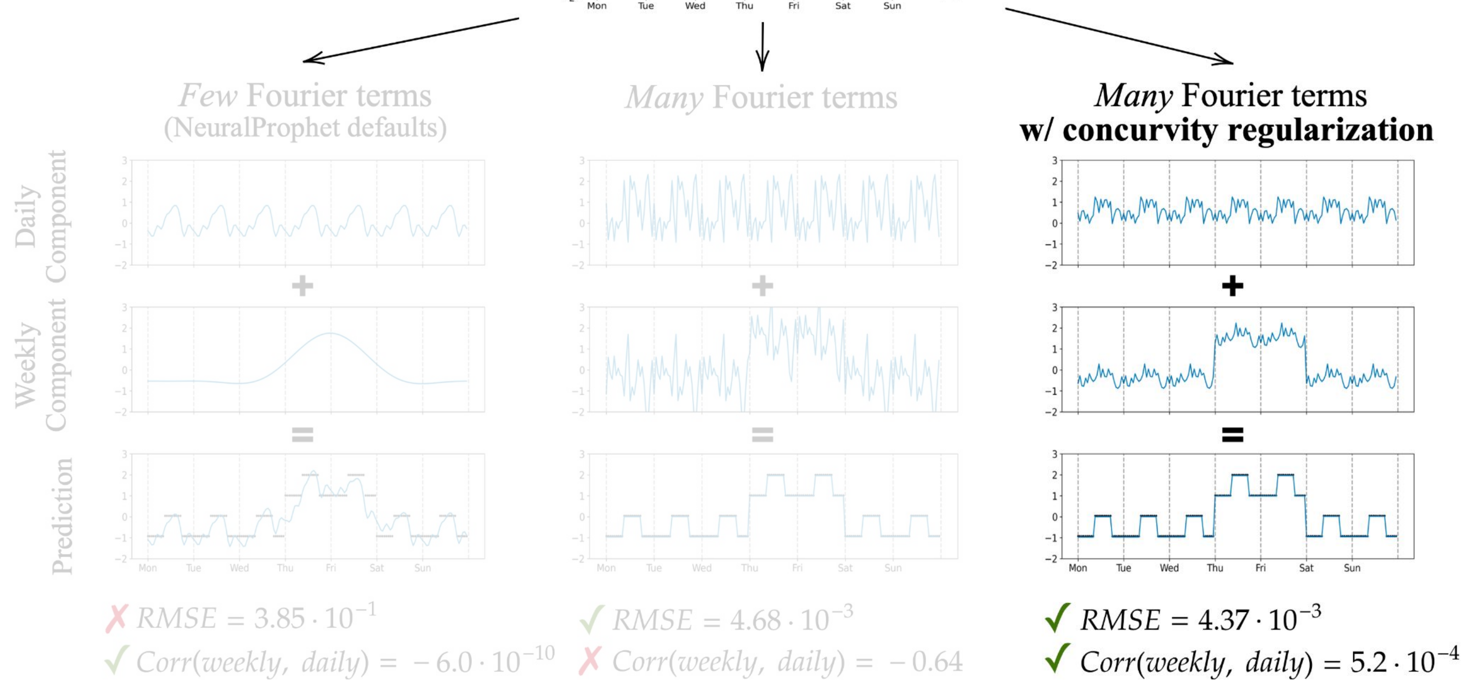
$$f(t) = S_{24h}(t) + S_{7d}(t)$$

3. Multicollinearity and Concurvity

Task: Fit periodic timeseries with daily and weekly seasonality



$$f(t) = S_{24h}(t) + S_{7d}(t)$$



4. Concurvity Regularizer

$$\min_{(f_1,\ldots,f_p)\in\mathcal{H}} \frac{1}{N}\sum_{l=1}^{N} L\Big(Y, \beta + \sum_{i=1}^{p} f_i(X_i)\Big) + \lambda \cdot R_\perp(\{f_i\}_i, \{X_i\}_i)$$

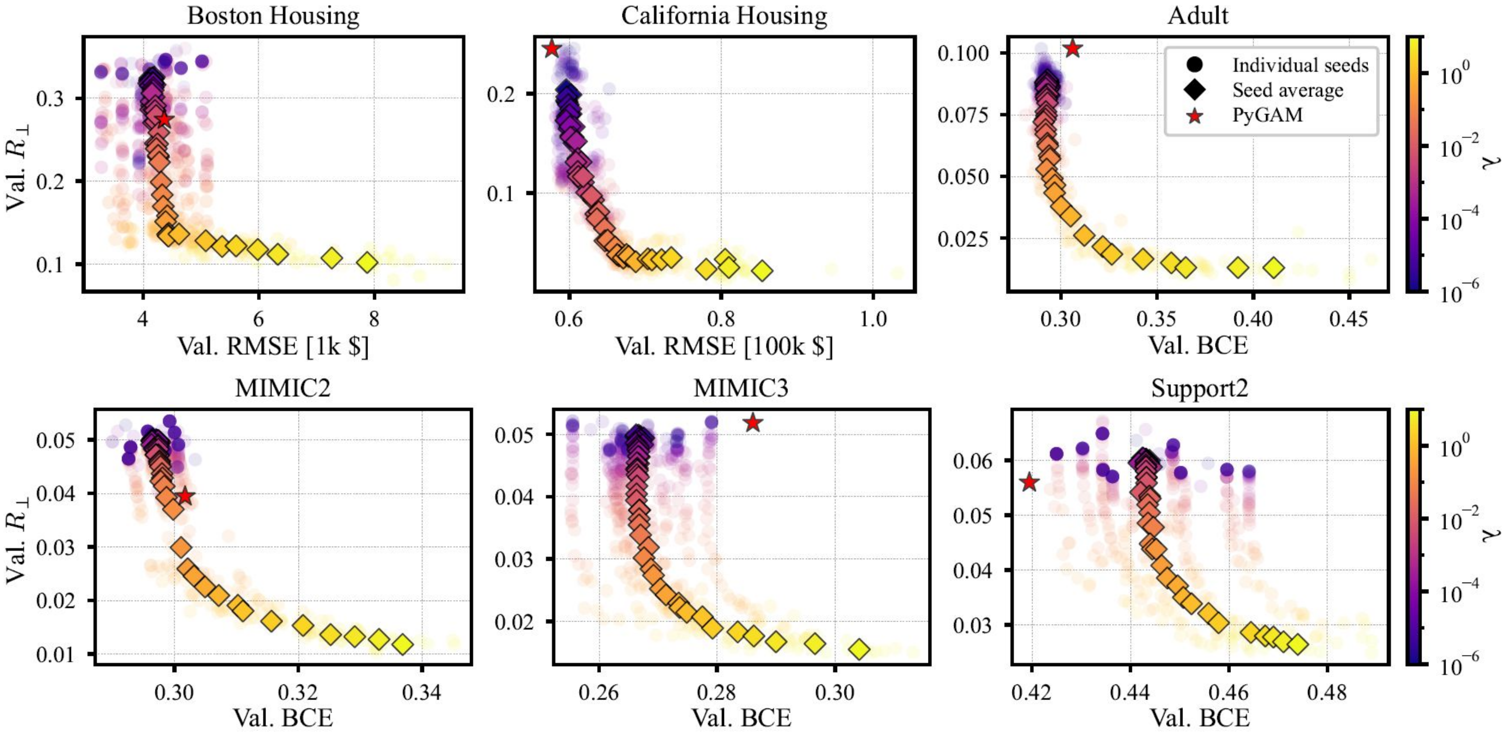
Definition: Concurvity Regularizer

$$R_\perp(\{f_i\}_i, \{X_i\}_i) := \frac{1}{p(p-1)/2}\sum_{i=1}^{p}\sum_{j=i+1}^{p} \left|\mathrm{Corr}\big(f_i(X_i), f_j(X_j)\big)\right|$$

Advantages of this approach:

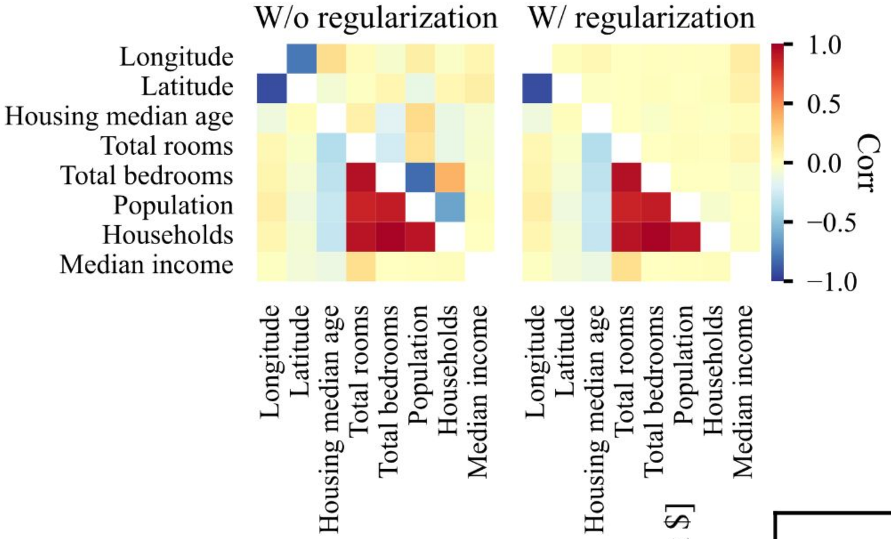
- Simple
- General
- Plug and Play

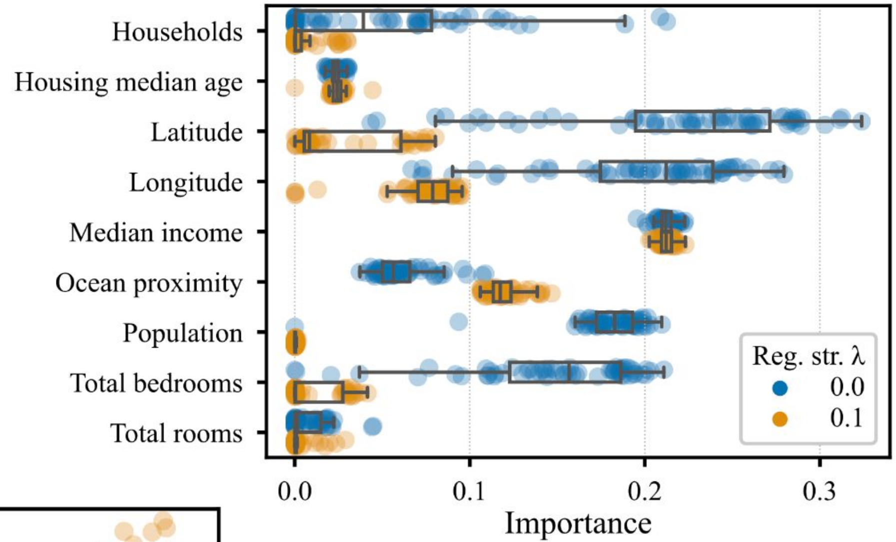
5. Results

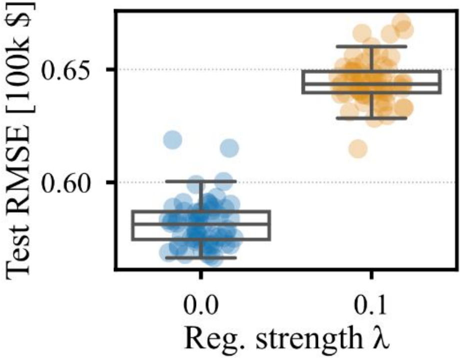


Finding: Concurvity can be reduced significantly without impacting model fit.

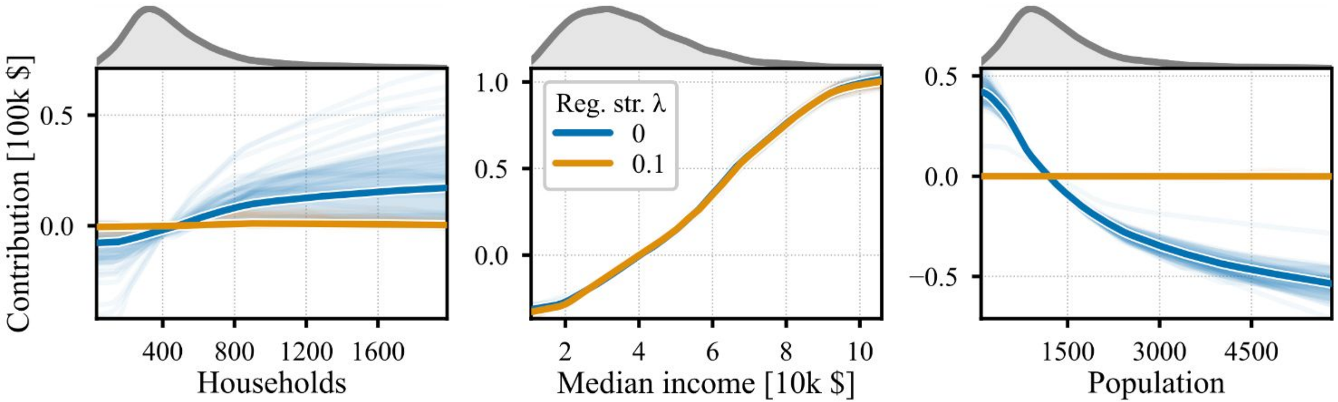
Case Study: California Housing



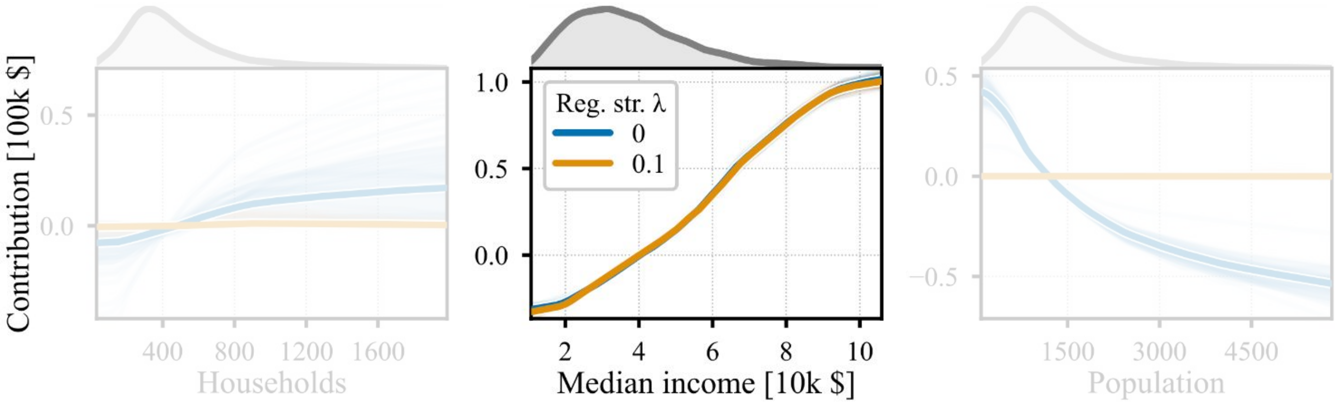




Case Study: California Housing



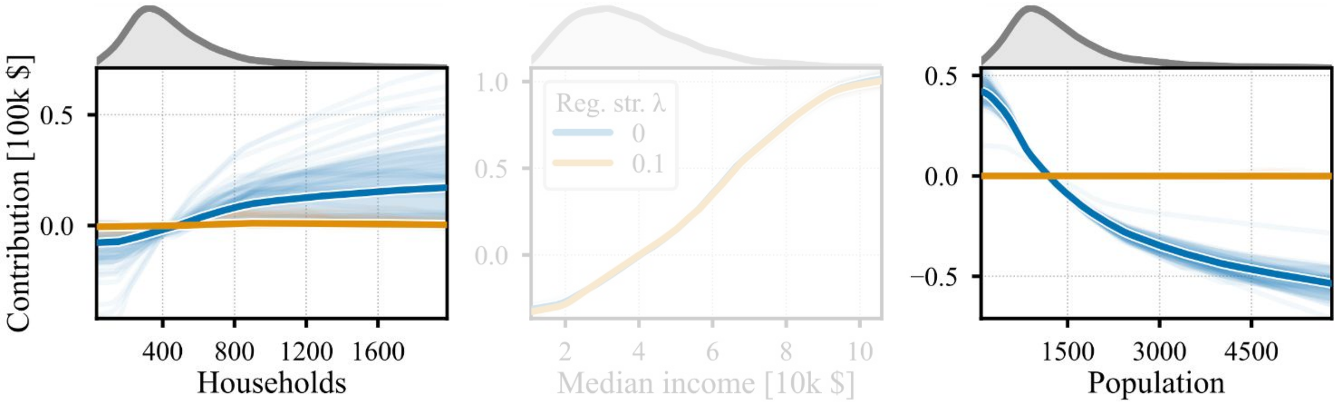
Case Study: California Housing



Finding:

- Uncorrelated features remain unaffected.

Case Study: California Housing



Finding:

- Correlated features often get almost pruned.

Key take home messages:

- We proposed a regularizer to mitigate concurvity in differentiable generalized additive models.
- We demonstrated its effectiveness at reducing concurvity while retaining model fit quality.

- Why ‘Curve Your Enthusiasm’? Watch out for concurvity to avoid drawing false conclusions from shape functions which hide concurvity.